Support is found at page 4, lines 19-20. These amendments address minor grammatical and 35 U.S.C. §112, second paragraph, issues and do not alter the scope of the pending claims.

The amendments to the specification eliminate reference to various internet URLs. The substituted addresses are well known and available to the public prior to the filing date of this application as evidenced by the web-site publications cited in the specification as filed. Other amendments to the specification are to correct grammatical/typographical errors and to comply with the requirements for sequence listings.

Allowed Subject Matter.

Applicants note with appreciation the Examiner's indication that claims 17-26, and 29-44 are allowed.

Information Disclosure Statement.

Applicants note that an Information Disclosure Statement (Form 1449 identifying references 1-94) submitted on August 11, 2000 **was not** considered and made of record in the present case. Applicants note that he subject Information Disclosure Statement was "hand-delivered" to Group 1600 on Friday August 11, 2000 at 2:41pm (see attached copy of postcard).

Applicants respectfully request that this Information Disclosure Statement, and the references cited therein, be expressly considered during the prosecution of this application and the references be made of record therein and appear among the "references cited" on any patent to issue therefrom.

In addition, Applicants provide herewith a supplemental Information Disclosure Statement and PTO-1449 form and request that this Supplemental Information Disclosure Statement, and the references cited therein, be expressly considered during the prosecution of this application and the references be made of record therein and appear among the "references cited" on any patent to issue therefrom.

Sequence Listing Rules.

The Examiner indicated that the application is not in compliance with sequence rules, 37 C.F.R. §§ 1.821-1.825. In particular, the examiner noted that sequences falling within the definitions set forth by the rule are found at page 10, line 9, and page 13, line 30. A disk containing the referenced sequence (SEQ ID NO:1) in computer readable form, and a paper copy of the

sequence information which has been printed from the floppy disk are provided herewith. The information contained in the computer readable disk was prepared through the use of the software program "PatentIn" and is identical to that of the paper copy.

Objection to the Specification.

The specification was objected to because it made reference to URLs on the internet. In particular, the Examiner objected to the internet references at page 24, line 14 and lines 1-12 on page 37. Pages 24 and 37 are amended with entry of this amendment to eliminate these references thereby obviating this objection.

Claim Objections.

Claims 1,and 15-16 were objected to because claim 1 recited a grammatically incorrect phrase "encoding two or more **a** biological molecules into character strings. . . ", and because claims 15 and 16 contained a period in the middle of the claim. Claims 1, 15, and 16 are amended herein to correct these issues thereby obviating these objections. It is noted that these amendments are grammatical in nature and **do not** alter the scope of the claims.

35 U.S.C. §112, Second Paragraph.

A) Use of the term "subunit".

Claims 1-16 were rejected under 35 U.S.C. §112, second paragraph, as allegedly indefinite because of the use of the term "subunits". In particular, the Examiner alleged that it was "unclear whether the subunits mean each letter corresponding to the nucleotides of a nucleic acid or a length of bytes required for the character string." Applicants respectfully traverse.

The term "subunit" is defined on page 8, line 32 through page 9, line 2 which states that:

> The term "subunit" **when used in reference to a biological molecule**
> refers to the characteristic "monomer" of which a biological is composed.
> Thus, for example, the subunit of a nucleic acid is a nucleotide, the subunit
> of a polypeptide is an amino acid, the subunit of a polysaccharide is a
> sugar, *etc.* [emphasis added]

In addition, claim 1 expressly recites:

> i) encoding two or more a biological molecules into character strings to
> provide a collection of two or more different initial character strings
> wherein **each of said biological molecules comprises at least about 10
> subunits**; [emphasis added]

Claim 1 clearly uses the term subunit in reference to a biological molecule. Thus, the term "subunit", as used in claim 1, refers to the characteristic "monomer" of which the biological is composed. Claim 1 is not indefinite in the use of the term "subunit" and the rejection of claims 1-16 under 35 U.S.C. §112, second paragraph, should be withdrawn.

B) Use of the phrase "30% sequence identity".

Claim 6 was rejected under 35 U.S.C. §112, second paragraph, as allegedly indefinite in the use of the phrase "biological molecules have at least 30% sequence identity". The Examiner alleged that is unclear to what the biological molecules are compared in determining the sequence identity. Applicants have amended claim 6 to recite that sequence identity is calculated between the biological molecules, *i.e.,* "the biological molecules have at least 30% sequence identity with each other."

C) Antecedent basis for "said method.

Claims 27 and 28 were rejected under 35 U.S.C. §112, second paragraph, because there allegedly was no antecedent basis for the term "said method". These claims are amended to eliminate this language thereby obviating this rejection. It is noted that these amendments do not alter the scope of the claims.

In view of the foregoing, Applicants submit that all of the rejections under 35 U.S.C. §112, second paragraph, should be withdrawn and that all claims now pending in this application are in condition for allowance. The issuance of a formal Notice of Allowance at an early date is respectfully requested.

If a telephone conference would expedite prosecution of this application, the Examiner is invited to telephone the undersigned at (510) 337-7871.

LAW OFFICES OF JONATHAN ALAN QUINE
P.O. BOX 458
Alameda, CA 94501
Tel: 510 337-7871
Fax: 510 337-7877

Respectfully submitted,

Tom Hunter
Reg. No: 38,498

APPENDIX A
"MARKED UP" CLAIMS ILLUSTRATING THE AMENDMENTS MADE TO THE
CLAIMS OF 09/495,668 WITH ENTRY OF THIS AMENDMENT

1. (Once amended) A method of populating a data structure with a plurality of character strings, said method comprising:

i) encoding two or more [a] biological molecules into character strings to provide a collection of two or more different initial character strings wherein each of said biological molecules comprises at least about 10 subunits;

ii) selecting at least two substrings from said character strings;

iii) concatenating said substrings to form one or more product strings about the same length as one or more of the initial character strings;

iv) adding the product strings to a collection of strings; and

v) optionally repeating steps (i) or (ii) through (iv) using one or more of said product strings as an initial string in the collection of initial character strings.

6. (Once amended) The method of claim 1, wherein said biological molecules have at least 30% sequence identity with each other.

15. (Once amended) The method of claim 1, wherein said coding, selecting, or concatenating is performed on a server. [The method of claim 1, wherein said coding, selecting, or concatenating is performed on a server.]

16. (Once amended) [The method of claim 1, wherein said coding, selecting, or concatenating is performed on a server.] The method of claim 1, wherein said coding, selecting, or concatenating is performed on a client linked to a network.[.]

27. (Once amended) The program of claim 17, wherein said [method further comprises]code additionally randomly [altering]alters one or more characters of said character strings.

28. (Once amended) The program of claim 27, wherein said [method further comprises]code additionally randomly [selecting and altering]selects and alters one or more occurrences of a particular preselected character in said character strings.

APPENDIX C

CLAIMS PENDING IN USSN 09/495,668 WITH ENTRY OF THIS AMENDMENT

1. A method of populating a data structure with a plurality of character strings, said method comprising:

i) encoding two or more biological molecules into character strings to provide a collection of two or more different initial character strings wherein each of said biological molecules comprises at least about 10 subunits;

ii) selecting at least two substrings from said character strings;

iii) concatenating said substrings to form one or more product strings about the same length as one or more of the initial character strings;

iv) adding the product strings to a collection of strings; and

v) optionally repeating steps (i) or (ii) through (iv) using one or more of said product strings as an initial string in the collection of initial character strings.

2. The method of claim 1, wherein said encoding comprises encoding one or more nucleic acid sequences into said character strings.

3. The method of claim 2, wherein said one or more nucleic acid sequences comprise a nucleic acid sequence encoding a known protein.

4. The method of claim 1, wherein said encoding comprises encoding one or more amino acid sequences into said character strings.

5. The method of claim 4, wherein said one or more amino acid sequences comprise a nucleic acid sequence encoding a known protein.

6. The method of claim 1, wherein said biological molecules have at least 30% sequence identity with each other.

7. The method of claim 1, wherein said selecting comprises selecting substrings such that the ends of said substrings occur in string regions of about 3 to about 20 characters that have higher sequence identity with the corresponding region of another of said initial character strings than the overall sequence identity between the same two strings.

8. The method of claim 1, wherein said selecting comprises selecting substrings such that the ends of said substrings occur in predefined motifs of about 4 to about 8 characters.

9. The method of claim 1, wherein said selecting and concatenating comprises concatenating substrings from two different initial strings such that the concatenation occurs in a region of about three to about twenty characters having higher sequence identity between said two different initial strings than the overall sequence identity between said two different initial strings.

10. The method of claim 1, wherein said selecting comprises aligning two or more of said initial character strings to maximize pairwise identity between two or more substrings of the character strings, and selecting a character that is a member of an aligned pair for the end of one substring.

11.   The method of claim 1, wherein said product strings are added to the collection only if they have greater than 30% sequence identity with the initial strings.

12.   The method of claim 1, wherein said method further comprises randomly altering one or more characters of said character strings.

13.   The method of claim 12, wherein said method further comprises randomly selecting and altering one or more occurrences of a particular preselected character in said character strings.

14.   The method of claim 1, wherein said coding, selecting, or concatenating is performed on an internet site.

15.   The method of claim 1, wherein said coding, selecting, or concatenating is performed on a server.

16.   The method of claim 1, wherein said coding, selecting, or concatenating is performed on a client linked to a network.

17.   A computer program product comprising computer code that
       i) encodes two or more a biological molecules into character strings to provide a collection of two or more different initial character strings wherein each of said biological molecules comprises at least about ten subunits;
       ii) selects at least two substrings from said character strings;
       iii) concatenates said substrings to form one or more product strings about the same length as one or more of the initial character strings;
       iv) adds the product strings to a collection of strings; and
       v) optionally repeats steps (i) or (ii) through (iv) using one or more of said product strings as an initial string in the collection of initial character strings.

18.   The program of claim 17, wherein said two or more biological molecules are nucleic acid sequences.

19.   The program of claim 17, wherein said two or more biological molecules are nucleic acid sequences of known proteins.

20.   The program of claim 17, wherein said two or more biological molecules are amino acid sequences

21.   The program of claim 17, wherein said biological molecules have at least 30% sequence identity.

22.   The program of claim 17, wherein said code selects substrings such that the ends of said substrings occur in string regions of about three to about twenty characters that have higher sequence identity with the corresponding region of another of said initial character strings than the overall sequence identity between the same two strings.

23. The program of claim 17, wherein said code selects substrings such that the ends of said substrings occur in predefined motifs of about 4 to about 8 characters.

24. The program of claim 17, wherein said code selects and concatenates substrings from two different initial strings such that the concatenation occurs in a region of about three to about twenty characters having higher sequence identity between said two different initial strings than the overall sequence identity between said two different initial strings.

25. The program of claim 17, wherein code selects substrings by aligning two or more of said initial character strings to maximize pairwise identity between two or more substrings of the character strings, and selecting a character that is a member of an aligned pair for the end of one substring.

26. The program of claim 17, wherein said product strings are added to the collection only if they have greater than 30% identity with the initial strings.

27. The program of claim 17, wherein said code additionally randomly alters one or more characters of said character strings.

28. The program of claim 27, wherein said code additionally randomly selects and alters one or more occurrences of a particular preselected character in said character strings.

27. The program of claim 17, wherein said method further comprises randomly altering one or more characters of said character strings.

28. The program of claim 27, wherein said method further comprises randomly selecting and altering one or more occurrences of a particular preselected character in said character strings.

29. The program claim 17, wherein said code is stored on media selected from the group consisting of magnetic media, optical media, optomagnetic media.

30. The program claim 17, wherein said code is in dynamic or static memory of a computer.

31. A label generating system for creating a plurality of related labels, said labeling system comprising:
>an encoder for encoding two or more initial strings from biological molecules;
>an isolator for identifying and selecting substrings from said two or more strings;
>a concatenator for concatenating said substrings;
>a data structure for storing the concatenated substrings as a collection of strings;
>a comparator for measuring the number and variability of the collection of strings and determining that sufficient strings exist in the collection of strings; and
>a command writer for writing the collection of strings into a raw string file.

32.   The system of 31, wherein said isolator comprises a comparator for aligning and determining regions of identity between said two or more initial strings;

33.   The system of 31, wherein said encoder comprises a means for encoding a nucleic acid sequence into a character string.

34.   The system of 31, wherein said encoder comprises a means for encoding an amino acid sequence into a character string.

35.   The system of claim 31, wherein said comparator comprises a means for calculating sequence identity.

36.   The system of claim 31, wherein said isolator selects substrings such that the ends of said substrings occur in string regions of about three to about 100 characters that have higher sequence identity with the corresponding region of another of said initial character strings than the overall sequence identity between the same two strings.

37.   The system of claim 31, wherein said isolator selects substrings such that the ends of said substrings occur in predefined motifs of about 4 to about 8 characters.

38.   The system of claim 31, wherein said isolator and concatenator individually or in combination concatenate substrings from two different initial strings such that the concatenation occurs in a region of about three to about 100 characters having higher sequence identity between said two different initial strings than the overall sequence identity between said two different initial strings.

39.   The system of claim 31, wherein said isolator aligns two or more of said initial character strings to maximize pairwise identity between two or more substrings of the character strings, and selecting a character that is a member of an aligned pair for the end of one substring.

40.   The system of claim 31, wherein said comparator adds strings to said data structure only if they have greater than 30% identity with the initial strings.

41.   The system of claim 31, further comprising an operator to randomly altering one or more characters of the character strings.

42.   The system of claim 41, wherein said operator randomly selects and alters one or more occurrences of a particular preselected character in said character strings.

43.   The system of claim 31, wherein data structure is a data structure that stores encoded nucleic acid sequences.

44.   The system of claim 31, wherein data structure is a data structure that stores encoded amino acid sequences.